AMERICAN PSYCHOLOGICAL ASSOCIATION
Journal of Experimental Psychology: General

© 2025 American Psychological Association
ISSN: 0096-3445

2025, Vol. 154, No. 10, 2926–2934
https://doi.org/10.1037/xge0001811

BRIEF REPORT

Executive Control Can Query Hidden Human Memories

Chong Zhao[1, 2], Keisuke Fukuda[3], and Geoffrey F. Woodman[4]
[1] Department of Psychology, University of Chicago
[2] Institute for Mind and Biology, University of Chicago
[3] Department of Psychology, University of Toronto Mississauga
[4] Department of Psychology, Vanderbilt Vision Research Center, Vanderbilt University

When we try to retrieve a representation from visual long-term memory, there is a chance that we will fail to recall seeing it even though the memory is stored in our brain. Here we show that although mechanisms of explicit memory retrieval are sometimes unable to retrieve stored memories, mechanisms of executive control can quickly query memory and determine if a representation is stored therein. Our findings suggest that the representations stored in human memory that cannot be accessed explicitly at that moment are nonetheless directly accessible by the brain's higher level control mechanisms.

Public Significance Statement
Scientists believe that our brains contain memories that we cannot access at that moment but that are still stored in the tissue. This study shows that certain circuits in our brains can determine that those memories are present, even when we cannot remember them. An international group of researchers recorded electrical brain activity from human subjects while they had their memories tested. The subjects' brain activity showed that they had stored a memory of an object, even when they pressed the button to report that they did not remember seeing that object previously in the experiment. Using these signals, researchers can infer what is stored in the inaccessible parts of human memory, as well as feed these signals into computer systems to accelerate learning and the adaptive control of machines.

Keywords: event-related potentials, visual memory, executive control, error detection, error-related negativity

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

We humans frequently forget that we have previously encountered important things in our environment (Wixted & Carpenter, 2007; Wixted, 2022), with forgetting becoming more pronounced during aging and clinical disorders (García-Martínez et al., 2023). As cognitive psychologists, we often use recognition memory tests to access the quality of human memory in the laboratory (Brown & Aggleton, 2001). These tasks generally involve presenting to-be-remembered stimuli, such as objects, with later testing requiring subjects to report whether an item is old or new, and their confidence in that behavioral report (Figure 1A). Although visual long-term memory is quite accurate (Brady et al., 2008; Shoval et al., 2023), it is believed that human memory stores more information than can be explicitly retrieved at any given moment in time (Rugg et al., 1998; Voss & Paller, 2009). Here we tested the hypothesis that those memories that

---

This article was published Online First July 21, 2025.
Rachel Wu served as action editor.
Geoffrey F. Woodman https://orcid.org/0000-0003-0946-9297

The data, stimuli, and experimental code are posted publicly on the Open Science Framework at https://osf.io/pk4nt/. This study was not preregistered. The work was presented at the Vision Science Society as a poster, and the poster abstract was posted at https://jov.arvojournals.org/article.aspx?articleid=2801291. The authors previously presented a version of these findings at the 2024 Vision Sciences Society meeting. The data and idea were not shared on any online website prior to the submission of the article. The authors have no conflicts of interest to declare that are relevant to the content of this article.

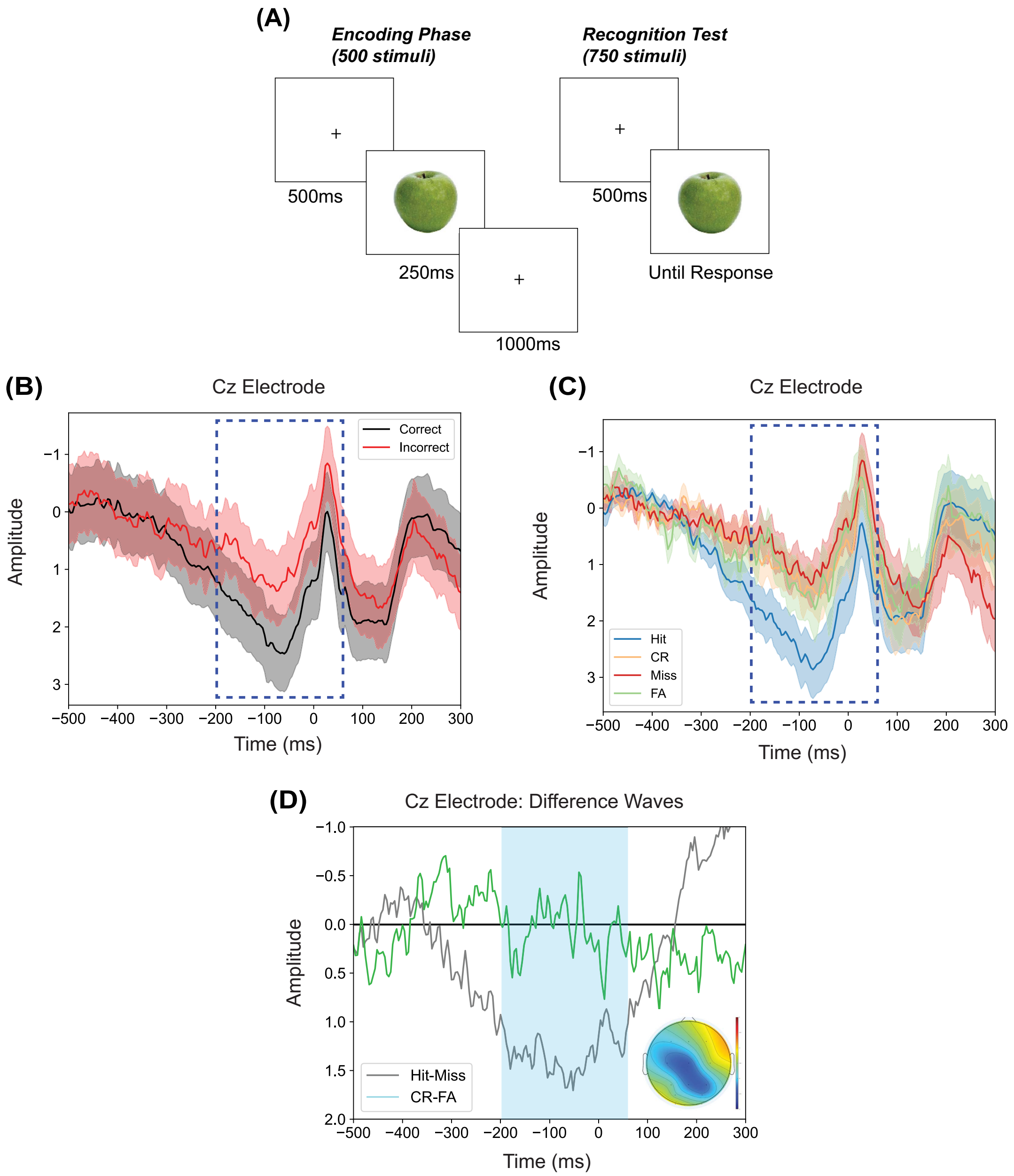
This research was supported by funding from the Directorate for Social, Behavioral and Economic Sciences, National Science Foundation (Grant BCS-2147064) and the National Institutes of Health (Grants P30-EY08126 and T32-EY007135) awarded to Geoffrey F. Woodman and Division of Behavioral and Cognitive Sciences and the Natural Sciences and Engineering Research Council (Grant 5009170) awarded to Keisuke Fukuda. The authors thank Gordon Logan, Sisi Wang, and Sean Polyn for the formative discussions of this work.

Chong Zhao played a lead role in formal analysis, investigation, and methodology and an equal role in writing–review and editing. Keisuke Fukuda played a lead role in data curation and an equal role in funding acquisition, project administration, and writing–review and editing. Geoffrey F. Woodman played a lead role in conceptualization, supervision, and writing–original draft and an equal role in funding acquisition and project administration.

Correspondence concerning this article should be addressed to Geoffrey F. Woodman, Department of Psychology, Vanderbilt Vision Research Center, Vanderbilt University, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240-7817, United States. Email: geoffrey.f.woodman@vanderbilt.edu

Figure 1

The Task Our Human Observers (N = 50) Performed and the Results From the Electroencephalogram Recordings



Note. Panel A: An example of the stimuli presented during the encoding and then the test phase of the experiment. Panel B: The waveforms time locked to button press onset for correct and incorrect trials plotted separately. Panel C: The waveforms sorted by whether the test image was of an old item shown during encoding (hit and miss) or a new item (CR or FA). Panel D: The difference waves and scalp distributions of the memory ERN (measuring in the blue region). ERN = error-related negativity; CR = correct rejection; FA = false alarm. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

cannot be explicitly reported in memory experiments can nonetheless be accessed by executive control mechanisms that monitor our behavioral response output and generate distinct electrophysiological signatures.

By recording humans' electroencephalogram (EEG), researchers have used the error-related negativity (or ERN) extensively to study executive control because it is elicited when humans make an error in many different settings (Dehaene et al., 1994; Gehring & Willoughby, 2002; Gehring et al., 2012). For the brain to detect this error, it must compare the behavioral response that was just made to the intended response for that trial. For perceptual-motor tasks, it is assumed that the brain codes the intended, correct response

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

(Crapse & Sommer, 2008; Nieuwenhuis et al., 2001). In the context of a memory experiment, the observation of an ERN has more revealing implications. That is, to elicit this activity associated with error detection following a memory error would show that executive control mechanisms of the brain had access to the necessary memory. The ERN can only be generated if the brain knows whether the test object was previously shown in the memory set. Therefore, we based our argument on the reverse inference that an existence of an ERN suggested error detection in memory tasks. To test this hypothesis, we accumulated a large data set from 50 individuals sampled across two institutions performing the same visual recognition memory task. This large data set afforded further comparisons of the error-related activity as a function of response speed and confidence.

Method

Transparency and Openness

This study was not preregistered before data analysis. We report how we determined all data exclusions, all manipulations, and all measures in the study. Experiments were programmed in MATLAB 2015. Analyses were performed in Python 3.7, with the package matplotlib, scipy, numpy, and seaborn for plotting. Data and materials used in this study are accessible to the public on the Open Science Framework repository (https://osf.io/pk4nt/).

Participants

We conducted a power analysis to determine the sample size needed for adequate statistical power, employing an effect size of 0.5 and a significance level of 0.05. The analysis revealed that a minimum of 34 subjects would be necessary to achieve a power of 0.8, while a sample size of 44 subjects would be required to attain a power of 0.9. Subjects were recruited from Vanderbilt University and the University of Toronto and paid for their participation. For the Vanderbilt sample, 13 participants self-reported as women and 10 as men (aged 18–32) took part after providing informed consent, with all procedures approved by Vanderbilt University's Institutional Review Board. Participants were compensated with $30. All participants self-reported being neurologically healthy, with normal or corrected-to-normal vision, and no color vision deficiencies. Data from three individuals were excluded due to incomplete session participation. We excluded seven participants due to excessive blinks (detected by vertical electrooculograms) during the response phase. After artifact rejection, 19 subjects were kept in our final analyses.

For the University of Toronto sample, 22 participants self-reported as women and 16 as men ($M = 20.32$ years old, $SD = 2.55$, range = 18–32) from the University of Toronto Mississauga community and they participated in return for monetary compensation (12 CAD/hr). Every participant provided written informed consent to the protocol approved by the Research Ethics Boards of the University of Toronto. Data obtained from seven participants were rejected due to excessive EEG artifacts (>40% trials rejected due to artifacts), resulting in 31 participants in the initial analysis. In total, our experiment contained 50 subjects for our final analysis, all of whom performed above chance on the memory task.

Stimuli and Procedure

The stimuli and task setup are depicted in Figure 1A. The stimuli were adapted from an established set of photographs (Brady et al., 2008). During the encoding task, participants were shown 500 sequential images of real-world objects, with brief pauses every 50 images. Participants were instructed to study each image while maintaining central fixation, preparing for a subsequent recognition memory test. Each trial began with a button press on a gamepad, followed by a 1,250 ms pre-encoding interval, a 250 ms presentation of the image, and then a 1,000 ms blank screen during encoding. A central fixation dot then signaled the start of the next trial. After the encoding phase, participants' resting-state EEG activity was recorded for 15 min with eyes open and closed. Memory for the images was assessed using a recognition test.

The recognition memory test commenced with the appearance of a central fixation dot. Participants initiated each trial with a button press and were instructed to maintain central fixation without blinking throughout the trial. Following a 1,250 ms blank screen, an image of a real-world object was displayed at the center of the screen, with a mix of new and previously seen images presented in random order. After the 1,250 ms image presentation, a blue and a red dot appeared, each on opposite sides of the image. Participants then indicated whether they remembered seeing the image during the study phase. The red dot's position indicated which side of the gamepad to press if the image was remembered, while the blue dot indicated the buttons to press if it was not. Participants used a scale to indicate their confidence, pressing different buttons for 100%, 80%, or 60% confidence, with participants being told that 50% represents a pure guess in a two alternative-forced-choice task. The red and blue dot positions were randomized in each trial. After responding, participants had a self-paced break to rest and blink. They were tested on 500 previously seen images and 250 new images.

EEG Acquisition and Preprocessing

EEG recordings were made using a right-mastoid reference, with offline rereferencing to the average of both mastoids. Electrodes were placed at standard 10–20 system locations (Fz, Cz, Pz, F3, F4, C3, C4, P3, P4, PO3, PO4, O1, O2, T3, T4, T5, and T6) and at custom locations OL (halfway between O1 and T3) and OR (halfway between O2 and T4). Eye movements were tracked with electrodes placed 1 cm lateral to the external canthi for horizontal movements and beneath the right eye for vertical movements and blinks. Signals were amplified with a gain of 20,000 bandpass filtered between 0.01 and 100 Hz and digitized at 250 Hz. Trials with horizontal eye movements exceeding a 30 μV threshold or eye blinks exceeding a 75 μV threshold were excluded from further analysis.

To isolate oscillatory activity within the theta frequency range (5–8 Hz), we applied a zero-phase bandpass filter to the EEG data. The filter was designed using the Butterworth implementation provided in the SciPy library in Python (Virtanen et al., 2020). Specifically, we implemented a second-order (equivalent to fourth-order forward–backward) Butterworth bandpass filter with cutoff frequencies set at 5 Hz (lower bound) and 8 Hz (upper bound). To avoid phase distortions, the filter was applied in both the forward and reverse directions using the filtfilt function, ensuring zero-phase shift in the filtered signal. This procedure is consistent with best practices for frequency-based EEG analyses (Keil et al., 2022).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Following filtering, we computed theta power as the root-mean-square amplitude of the filtered signal within each trial epoch. We also visually inspected the filtered signal to ensure that the theta activity was not contaminated by high-frequency noise or low-frequency drift. The choice of the theta frequency band (5–8 Hz) was motivated by prior research linking theta oscillations to cognitive control and working memory processes (Cavanagh & Frank, 2014; Jensen & Tesche, 2002). All preprocessing and filtering steps were performed using Python (Version 3.6) with the SciPy library (Virtanen et al., 2020). We calculated the Bayes factor with the PyMC package (Python 3).

Event-Related Potential Analyses

To examine the event-related potentials associated with the memory test response, data were time locked to the button press that indicated the response during each trial, analyzing waveforms from −500 ms to +300 ms relative to the motor response. To observe the ERN emerge around the time of response, we baseline-corrected each EEG segment to the mean amplitude from −500 ms to −300 ms before the measurement epoch of interest. We entered the mean amplitudes measured from −200 to +50 ms around the time of the button press into an analysis of variance with the factors of correctness (correct vs. error) and electrode (Fz, Cz, vs. Pz), with subsequent analyses examining median splits based on each individual's median response time (RT) or confidence rating across all stimuli. We also analyzed our data for the error positivity (or Pe) which follows the ERN and is maximal over parietal electrodes. We did not find that the positivity following the ERN was significantly modulated by the status of the test object, unlike the ERN. In the Appendix, we show the FN400 waveforms. The FN400 or frontal old/new effect appears to measure the familiarity, implicit memory, or perceptual fluency with which a stimulus is processed (Voss & Federmeier, 2011). We show these waveforms to verify that the pattern of effects across the ERN is different than those of the FN400, which overlaps it in time.

Results

As shown in Figure 1B and 1C, we found that subjects' waveforms were more negative just as they were about to commit an error. The ERN is defined as a more negative potential around the time of a behavioral response on error trials relative to trials with correct responses, Fz: $t(49) = 2.38$, $p = .02$, Cohen's $d = 0.34$; Cz: $t(49) = 3.37$, $p = .001$, Cohen's $d = 0.48$. This effect began approximately 200 ms before the button press was made and continued for approximately 100 ms afterward. It is known that the ERN can precede the commission of the erroneous behavioral response, presumably because the difference between the actual processing state and the intended processing state is detected prior to the final stages of response execution (Falkenstein et al., 2000; Rodriguez-Fornells et al., 2002). The scalp distribution of this effect is similar to previous studies of the ERN using other tasks (Figure 1D blue inset; Dehaene et al., 1994; Reinhart et al., 2012). Inferential statistical tests showed that error trials elicited more negative potentials relative to correct trials evidenced by a main effect of accuracy, $F(1, 49) = 11.56$, $p = .001$, $\eta_p^2 = 0.13$, in the analysis of variance of voltage, and because this difference was larger in amplitude at central relative to frontal electrodes, we observed an interaction of Electrode × Trial Type, $F(1, 49) = 7.68$, $p = .008$, $\eta_p^2 = 0.1$, due to larger amplitude ERN at Cz than Fz.

Our next goal was to determine the relationship between the ERN and the nature of the behavioral responses that it is time locked to. The proportion of trials, mean RT, and confidence ratings measured across hit, miss, correct rejection, and false alarm trials are summarized in Table 1. Our first relevant observation is that the ERN exhibited a large amplitude difference between miss trials and hit trials, $t(49) = 2.98$, $p = .004$, Cohen's $d = 0.42$, but there was not a significant difference in the response-locked waveforms between false alarm and correct rejection trials, $t(49) = 0.08$, $p = .93$; see Figure 1C. This observation of an ERN elicited on miss trials relative to hit trials (i.e., when old test items are shown), but not false alarms relative to correct rejections (i.e., when new test items are shown), indicates that the executive control mechanism that generates the ERN can detect when a test item matches a representation in memory but does not return diagnostic evidence about the behavioral response being inaccurate when no memory match can be found. This result is similar to other cognitive tasks in which observers need to determine whether a given object is absent from a visual or memory set (Chun & Wolfe, 1996; Treisman & Gormican, 1988; Zhao et al., 2023). Next, we analyzed our data to determine whether this error-related activity was sensitive to the speed or confidence of the observers' judgments about whether they remembered something.

To determine whether the memory-related ERN was related to the speed of subjects' behavior, we analyzed the test stimulus-locked event-related potentials based on accuracy and divided by median response time; the ERN amplitude was found to be more negative for incorrect trials compared to correct trials, main effect of accuracy: $F(1, 49) = 4.75$, $p = .03$, $\eta_p^2 = 0.10$, and more negative for slower trials compared to faster ones, main effect of RT: $F(1, 49) = 5.42$, $p = .02$, $\eta_p^2 = 0.11$; see Figure 2A. In line with the results in Figure 1, fast hit and fast miss trials exhibited a significant difference in amplitude between each other, $t(49) = 2.57$, $p = .01$, Cohen's $d = 0.38$, whereas fast correct rejection and fast false alarm trials did not, $t(49) = 0.29$, $p = .77$, $BF_{10} = 0.11$; see Figure 2B. This suggests that a more negative ERN amplitude was associated with a more

Table 1

Summary of the Number of Trials, Mean Response Time, and Confidence Ratings for Hit, Miss, Correct Rejection, and False Alarm Trials

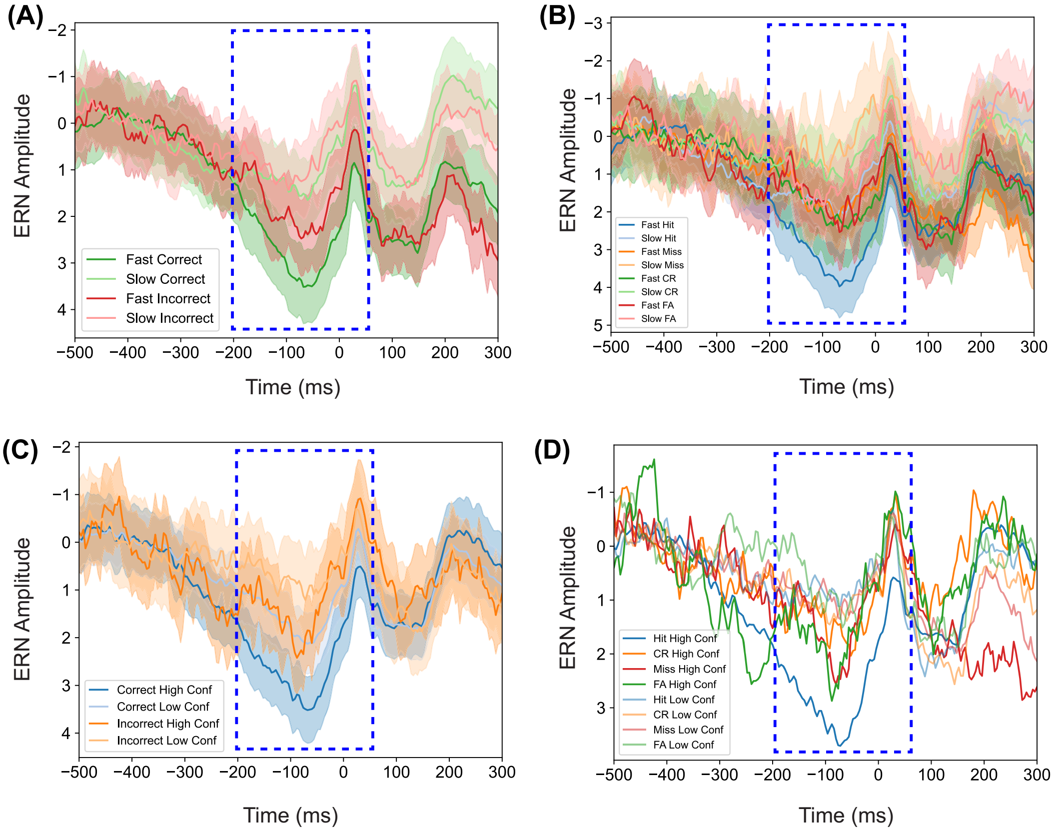
Response type	Proportion of trials	Mean RT (ms)	Mean confidence rating (60% = low, 80% = medium, 100% = high)
Hit	42.82%	567.05	90.19
Miss	24.07%	580.85	79.64
Correct rejection	23.41%	589.52	82.71
False alarm	9.70%	660.09	78.51

Note. RT = response time.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Figure 2

Results Showing Relationships Between Subjects' Behavior and the ERN



Note. Panel A: The ERN plotted as a function of accuracy and RT at electrode Cz. Panel B: The ERN as a function of response speed, using a median split. Panel C: The ERN as a function of response speed and the type of test item presented. Panel D: The ERN as a function of response type (hit, miss, CR, or FA) and confidence. ERN = error-related negativity; RT = response time; CR = correct rejection; FA = false alarm; Conf = confidence. See the online article for the color version of this figure.

challenging retrieval process during the test phase, resulting in longer response times.

Presumably, executive control mechanisms can query both explicitly available and currently unavailable representations from memory, such that the amplitude of this metric should be correlated with subjective reports of confidence, as with the RT analyses that we just discussed. These subjective confidence reports were typically collected using a binary measure of participants' sureness about their responses. Prior research with the Flanker task (e.g., Scheffers & Coles, 2000) found that the ERN was larger when participants were confident in their errors compared to when they were uncertain. This study also showed that this pattern extended to correct responses, with participants exhibiting a higher correct response negativity for confident correct responses than for uncertain ones. Thus, we tested for an effect of confidence using a preplanned *t* test.

In Figures 2C and 2D, we show the waveforms as a function of confidence in visual memory judgment, with both the correct and error waveforms changing with confidence. The analysis of variance exhibited a significant effect of correctness, $F(1, 49) = 7.44$, $p = .009$, $\eta_p^2 = 0.13$, and a significant interaction of electrode and confidence, $F(1, 49) = 5.69$, $p = .02$, $\eta_p^2 = 0.10$, due to the modulation by confidence being maximal at location Cz for correct responses. The Confidence × Accuracy interaction was not significant, $F(1, 49) = 0.62$, $p = .43$, $\eta_p^2 = 0.01$; however, in our planned comparison, we found a significant effect of confidence, correct high confidence being higher than correct low confidence at Cz, $t(49) = 3.88$, $p = .003$; incorrect high versus low confidence waveforms were not different at Cz, $t(49) = 0.49$, $p = .63$, the electrode shown in Figure 2A–2D. These findings suggest that the pattern of ERN activity was sensitive to the confidence that subjects had when they made their behavioral responses, as we would expect if this mechanism of executive control can access all the contents of memory.

To determine the generality of our findings, we conducted analyses on another neural signature of cognitive control, the midline theta power (Cavanagh & Frank, 2014). In simple motor judgments, the midline theta power (5–8 Hz) magnitude measured shortly after a motor response is higher for erroneous responses than correct responses. Prior studies on midline theta had used a wide range of tasks, such as sustained attention to cue tasks (van Driel et al., 2012; theta power peaking 50 ms–300 ms postresponse) and response conflict tasks (Valadez & Simons, 2018; theta power peaking 0 ms–400 ms postresponse). Therefore, we calculated theta power in 50-ms rolling windows from 0 ms up to 400 ms after the response. Here, we also found baselined theta power was higher immediately following

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

incorrect responses than correct responses, measured in the window from 150 ms to 250 ms after memory responses were made, $t(49) = 3.56$, $p < .001$ (Figure 3A). However, different from the ERN, the midline theta power seemingly reflected a more general error detection process, in that the error responses always exhibited a more negative theta power compared to correct responses regardless of trial type, hit versus miss, $t(49) = 3.16$, $p = .003$; correct rejection versus false alarm, $t(49) = 2.93$, $p = .005$; see Figure 3B. To test if the two effects were functionally independent from each other during recognition memory decision, we bandpass filtered our EEG signals to eliminate theta oscillations in our raw signals. The remaining signals, with no theta power in the time domain, replicated our findings with the unfiltered data, hit versus miss, $t(49) = 2.88$, $p = .006$; correct rejection versus false alarm, $t(49) = 0.06$, $p = .95$; see Figure 3C. Therefore, we conclude that the mechanism indexed by the ERN is sensitive to what is stored in memory during recognition judgments when participants studied the images during the encoding phase (i.e., hit vs. miss), while the midline theta power reflects a general error detection mechanism that is insensitive to what is held in visual long-term memory.

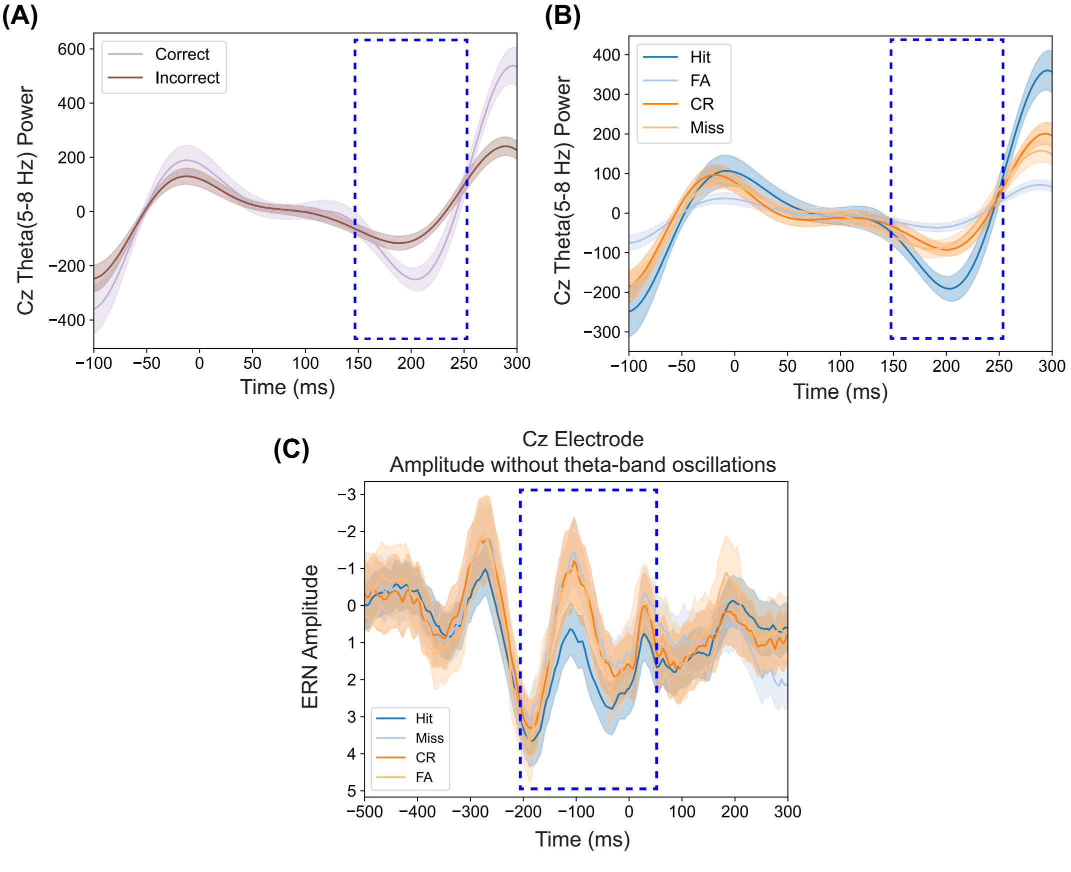
In classic studies of the ERN, researchers reported that a higher ERN amplitude led to more compensatory behaviors on the trial following the erroneous response (Gehring et al., 1993). In the present study, subjects reported when they were ready for another memory test by initiating the next test with a space bar press. Our findings appear inconclusive as yet, because we lacked sufficient resolution to observe a significant slowing of initiation time following a miss compared to a hit, $t(49) = 1.72$, $p = .09$, although such a trend was observed, with no such slowing between false alarms and correct rejections, $t(49) = 1.07$, $p = .29$, mirroring the pattern of medial-frontal negativities on these trials. We believe that it would be fruitful for subsequent studies to focus on determining whether error detection during testing memory results in compensatory behavior on the next event.

Discussion

We observed a medial-frontal negativity around the time of our subjects' button presses that was more negative for error responses

Figure 3

Plots Showing the Relationships Between Midline Theta (6–8 Hz) Power, Subjects' Behavior, and Error-Related Negativity



Note. Panel A: Theta power sorted by the accuracy on each trial. Panel B: Theta power sorted by whether the test image was of an old item shown during encoding (hit or miss) or a new item (correct rejection or false alarm). Panel C: The ERN, after filtering out theta band (6–8 Hz) oscillations, as a function of whether the test image was of an old item shown during encoding (hit or miss) or a new item (correct rejection or false alarm). The filtered ERN, with no theta band oscillations, replicated what we found in Figure 1C. ERN = error-related negativity; CR = correct rejection; FA = false alarm. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

than for correct responses during visual long-term memory recognition testing. This medial-frontal negativity had a scalp distribution similar to that observed in previous studies of the ERN and was related to the speed and confidence of those trial-by-trial recognition judgments.

From a neural perspective, previous research has shown that ERN amplitudes tend to be larger for reported errors than unreported errors across various tasks, including the Flanker task (Scheffers & Coles, 2000; Maier et al., 2008), visual search (Woodman, 2010), and the antisaccade task (Wessel et al., 2011). This supports the idea that larger ERN amplitudes are associated with increased awareness of errors. In line with this, studies using the Flanker task have found that ERN amplitudes are greater when participants are confident in their errors compared to when they are uncertain (Scheffers & Coles, 2000). Similarly, this pattern extends to correct responses, with higher correct response negativity observed for confident correct responses than for uncertain ones. These findings suggest that the ERN reflects participants' awareness of their response accuracy, rather than merely signaling error commission. Applying this to visual long-term memory task used in our study, we propose that participants were more aware of their errors when the images had been previously studied (hit vs. miss) than when they were new (correct rejection vs. false alarm). Although participants may have recognized their errors when making a miss response, this awareness does not necessarily prevent errors. Instead, it simply reflects the brain's detection of an erroneous response. While the ERN acts as a neural signal of error detection, it does not inherently correct behavior. Relatedly, prior research suggests that larger ERN amplitudes are linked to increased likelihood of error correction if given a second chance (Endrass et al., 2007; Wessel et al., 2011). Future research is needed to explore the role of ERN in error correction in visual long-term memory tasks.

Moreover, our findings align with the broader error-monitoring literature, which suggests that ERN amplitude scales with the degree of conflict or error awareness (Yeung et al., 2004). However, we propose that this awareness may vary between response types. Specifically, participants may have been more aware of their errors when they failed to recognize previously studied images (misses) than when they incorrectly identified new images as studied (false alarms). This aligns with the idea that continued processing of a stimulus following an error can heighten conflict detection, as reflected by increased ERN amplitudes (Yeung et al., 2004). By interpreting the ERN as an index of memory-based error detection, our findings contribute to the growing body of work suggesting that the ERN reflects not just motor-related errors but also cognitive conflict arising from discrepancies between expected and actual responses.

The present empirical observations have important theoretical and practical implications. Theoretically, the present findings show that high-level executive control mechanisms appear to be able to access the memory representations that are stored in the human brain, apparently via a route that is independent of the explicit retrieval mechanisms. Although it has been known that the human brain can implicitly access memory representations that cannot be explicitly retrieved by performing tasks with greater efficiency across experience (Reber, 2013; Roediger, 1990; Rugg et al., 1998), the present article shows for the first time that the brain has mechanisms that can quickly query those memories, even those memories that cannot be explicitly accessed to guide behavior at that moment. This counterintuitively suggests that executive control mechanisms may have memory access that even memory mechanisms lack. Given the importance of detecting performance failures in large-scale cognitive models, such as prediction errors in reinforcement learning and even in threat processing, future studies are needed to understand in the decision-making stage of long-term memory tasks despite its long focus (Wixted & Carpenter, 2007).

Practically, the present finding suggests that it may be possible to develop brain–computer interfaces that can detect when the user has a memory stored, even when they fail to report it. That is, in principle an algorithm trained on a person's own medial-frontal negativities could probe the depths of memory storage to determine whether that individual has a currently hidden memory. Clearly the ethical use of this kind of memory readout ability will require more extensive discussion in broader society.

Constraints on Generality

In our study, we collected data from healthy young adults. As such, we currently draw conclusions about how memories are stored and accessed in the healthy human brain. It is possible that these processes operate differently across individuals, as the ERN and other components are known to change across disease states (Reinhart et al., 2015).

References

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, *2*(1), 51–61. https://doi.org/10.1038/35049064

Cavanagh, J. F., & Frank, M. J. (2014, August). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology*, *30*(1), 39–78. https://doi.org/10.1006/cogp.1996.0002

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, *9*(8), 587–600. https://doi.org/10.1038/nrn2457

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*(5), 303–305. https://doi.org/10.1111/j.1467-9280.1994.tb00630.x

Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: Aware and unaware errors in an antisaccade task. *European Journal of Neuroscience*, *26*(6), 1714–1729. https://doi.org/10.1111/j.1460-9568.2007.05785.x

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components of reaction errors and their functional significance: A tutorial. *Biological Psychology*, *51*(2–3), 87–107. https://doi.org/10.1016/S0301-0511(99)00031-9

García-Martínez, M., Sánchez-Juan, P., & Butler, C. R. (2023). A review of accelerated long-term forgetting in Alzheimer's disease: Current situation and prospects. *Neuropsychology*, *37*(6), 673–682. https://doi.org/10.1037/neu0000827

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*(6), 385–390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error-related negativity (ERN/Ne). In S. J. Luck & E. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 231–291). Oxford University Press.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*(8), 1395–1399. https://doi.org/10.1046/j.1460-9568.2002.01975.x

Keil, A., Bernat, E. M., Cohen, M. X., Ding, M., Fabiani, M., Gratton, G., Kappenman, E. S., Maris, E., Mathewson, K. E., Ward, R. T., & Weisz, N. (2022, May). Recommendations and publication guidelines for studies using frequency domain and time-frequency domain analyses of neural time series. *Psychophysiology*, *59*(5), Article e14052. https://doi.org/10.1111/psyp.1405

Maier, M., Steinhauser, M., & Hübner, R. (2008, December). Is the error-related negativity amplitude related to error detectability? Evidence from effects of different error types. *Journal of Cognitive Neuroscience*, *20*(12), 2263–2273. https://doi.org/10.1162/jocn.2008.20159

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence for an antisaccade task. *Psychophysiology*, *38*(5), 752–760. https://doi.org/10.1111/1469-8986.3850752

Reber, P. J. (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research. *Neuropsychologia*, *51*(10), 2026–2042. https://doi.org/10.1016/j.neuropsychologia.2013.06.019

Reinhart, R. M. G., Carlisle, N. B., Kang, M. S., & Woodman, G. F. (2012). Event-related potentials elicited by errors during the stop-signal task. II: Human effector-specific error responses. *Journal of Neurophysiology*, *107*(10), 2794–2807. https://doi.org/10.1152/jn.00803.2011

Reinhart, R. M. G., Zhu, J., Park, S., & Woodman, G. F. (2015). Medial-frontal stimulation enhances learning in schizophrenia by restoring prediction-error signaling. *Journal of Neuroscience*, *35*(35), 12232–12240. https://doi.org/10.1523/JNEUROSCI.1717-15.2015

Rodriguez-Fornells, A., Kurzbuch, A. R., & Münte, T. F. (2002). Time course of error detection and correction in humans: Neurophysiological evidence. *The Journal of Neuroscience*, *22*(22), 9990–9996. https://doi.org/10.1523/JNEUROSCI.22-22-09990.2002

Roediger, H. L., III. (1990). Implicit memory: Retention without remembering. *American Psychologist*, *45*(9), 1043–1056. https://doi.org/10.1037/0003-066X.45.9.1043

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*(6676), 595–598. https://doi.org/10.1038/33396

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 141–151. https://doi.org/10.1037/0096-1523.26.1.141

Shoval, R., Gronau, N., & Makovski, T. (2023). Massive visual long-term memory is largely dependent on meaning. *Psychonomic Bulletin & Review*, *30*(2), 666–675. https://doi.org/10.3758/s13423-022-02193-y

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*(1), 15–48. https://doi.org/10.1037/0033-295X.95.1.15

Valadez, E. A., & Simons, R. F. (2018, April). The power of frontal midline theta and post-error slowing to predict performance recovery: Evidence for compensatory mechanisms. *Psychophysiology*, *55*(4), Article e13010. https://doi.org/10.1111/psyp.13010

van Driel, J., Ridderinkhof, K. R., & Cohen, M. X. (2012). Not all errors are alike: Theta and alpha EEG dynamics relate to differences in error-processing dynamics. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(47), 16795–16806. https://doi.org/10.1523/JNEUROSCI.0802-12.2012

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Voss, J. L., & Federmeier, K. D. (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology*, *48*(4), 532–546. https://doi.org/10.1111/j.1469-8986.2010.01085.x

Voss, J. L., & Paller, K. A. (2009). An electrophysiological signature of unconscious recognition memory. *Nature Neuroscience*, *12*(3), 349–355. https://doi.org/10.1038/nn.2260

Wessel, J. R., Danielmeier, C., & Ullsperger, M. (2011). Error awareness revisited: Accumulation of multimodal evidence from central and autonomic nervous systems. *Journal of Cognitive Neuroscience*, *23*, 3021–3036. https://doi.org/10.1162/jocn.2011.21635

Wixted, J. T. (2022). The enigma of forgetting. *Proceedings of the National Academy of Sciences*, *119*(12), Article e2201332119. https://doi.org/10.1073/pnas.2201332119

Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science*, *18*(2), 133–134. https://doi.org/10.1111/j.1467-9280.2007.01862.x

Woodman, G. F. (2010). Masked targets trigger event-related potentials indexing shifts of attention but not error detection. *Psychophysiology*, *47*, 410–414. https://doi.org/10.1111/j.1469-8986.2009.00948.x

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959. https://doi.org/10.1037/0033-295x.111.4.939

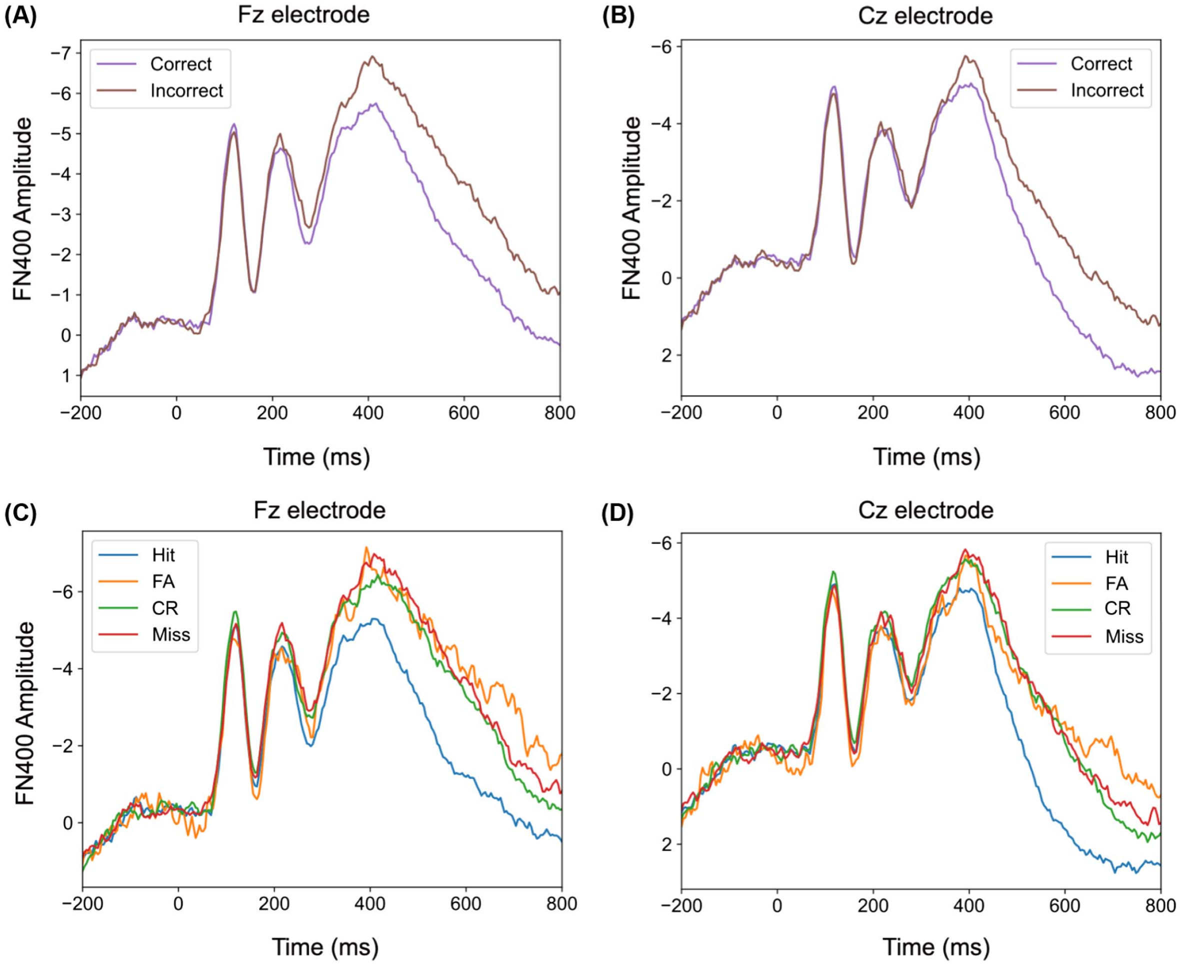
Zhao, C., Fukuda, K., & Woodman, G. F. (2023). Target recognition and lure rejection: Two sides of the same memorability coin? *Visual Cognition*, *31*(9), 633–641. https://doi.org/10.1080/13506285.2024.2315803

(*Appendix follows*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Appendix

FN400 Component From Electrodes Fz and Cz



Note. The analyses of additional ERP waveforms time locked to the test stimulus onset. Panel A: The FN400 or frontal old/new effect as a function of whether the subject would get the trial correct or not. Panel B: The scalp distribution of the FN400 was frontal, maximal at Fz. Panel C: The waveforms sorted by whether the test image was of an old item shown during encoding (hit and miss) or a new item (CR or FA). Panel D: The same waveforms shown at electrode Cz. FN = frontal negative; FA = false alarm; CR = correct rejection; ERP = event-related potential. See the online article for the color version of this figure.

Received October 2, 2024
Revision received May 15, 2025
Accepted May 22, 2025 ■